LENGTH AND SURFACE AREA ESTIMATION
UNDER SMOOTHNESS RESTRICTIONS

BEATRIZ PATEIRO-LÓPEZ * AND

ALBERTO RODRÍGUEZ-CASAL,* ** *Universidade de Santiago de Compostela*

Abstract

The problem of estimating the Minkowski content $L_0(G)$ of a body $G \subset \mathbb{R}^d$ is considered. For $d = 2$, the Minkowski content represents the boundary length of G. It is assumed that a ball of radius r can roll inside and outside the boundary of G. We use this shape restriction to propose a new estimator for $L_0(G)$. This estimator is based on the information provided by a random sample, taken on a square containing G, in which we know whether a sample point is in G or not. We obtain the almost sure convergence rate for the proposed estimator.

Keywords: Minkowski content; nonparametric set estimation; length estimation; convex set; r-convexity

2000 Mathematics Subject Classification: Primary 62G07
Secondary 62G20

1. Introduction

The aim of this work is to propose an estimator for the surface area of G. Therefore, in the two-dimensional case, $d = 2$, our goal is the estimation of the boundary length of G. This problem has been extensively considered, for example, in stereology; see [1, pp. 162–168] and the references therein.

The length and surface area estimation can be thought of as a further step in nonparametric set estimation theory. One of the goals in set estimation is to recover a set G, given a random sample of points whose distribution is related to G. The set of interest may be the support of a density, while the data is a sample drawn from this density. See [3] for more details. From this set estimation point of view, the estimation of the length and surface area of G can be seen as the problem of estimating an important geometric characteristic of a set which gives useful information about its shape. For example (for the bidimensional case), the 'contour index', which is defined as the ratio between the boundary length and the squared root of the area, provides a scale-invariant measurement of boundary roughness. Its minimal value $(2\sqrt{\pi})$ is attained by the circle and it increases as the set becomes more fragmented. The contour index has been used as an auxiliary diagnosis criterion in medical imaging. For example, in oncology the irregularity in the border of a tumor may suggest a bad prognosis because the damage is highly disseminated. See [4] for more details.

The case where G is assumed to be convex has been extensively analyzed in the literature, probably owing to the fact that in this case there exists a quite natural estimator of G: the convex hull of the sample. The perimeter and surface area of the convex hull of the sample can

be successfully used for estimating the length and surface area of G. See [2] and [9] for more details. However, assuming that G is convex may be too restrictive in practice. In a more general framework there is no obvious estimator to be used. Recently, in [4], a universally consistent estimator of the surface area $L_0(G)$ was proposed, without assuming any shape restriction on G (except that $L_0(G)$ exists). Unfortunately, this estimator achieves a relatively slow rate of convergence when G is smooth. The objective of this work is to propose an estimator of $L_0(G)$ which improves the convergence rate attained in [4] for a large family of smooth sets which includes convex sets. See Section 2 for the definition of this family.

The paper is organized as follows. In Section 2 we establish some notation and the precise definition of the 'surface area'. We also define the class of smooth sets we will work with. The sampling model and the proposed estimator is introduced in Section 3. Finally, in Section 4 the almost sure convergence rate for the proposed estimator is obtained.

2. Notation and previous concepts

Let G be a compact subset of Euclidean space $\mathbb{R}^d$. We assume, without loss of generality, that G is a subset of the open unit square $(0, 1)^d$.

We denote by $\| \cdot \|$ and $\langle \cdot, \cdot \rangle$ the Euclidean norm and the inner product of Euclidean space, respectively. The d-dimensional Lebesgue measure is denoted by μ. We respectively denote by $B(x, r)$ and $\mathring{B}(x, r)$ the closed and open ball with center x and radius r. In order to simplify the notation, we write B and $\mathring{B}$ to refer to $B(0, 1)$ and $\mathring{B}(0, 1)$, respectively.

If $A \subset \mathbb{R}^d$ then A^c, int(A), $\bar{A}$, and ∂A will stand for the complement, interior, closure, and boundary of A, respectively. The ε-neighborhood of A is defined by

$$B(A, \varepsilon) := \bigcup_{a \in A} B(a, \varepsilon) = \{x \in \mathbb{R}^d : d(x, A) \le \varepsilon\},$$

where $d(x, A) = \inf\{\|x - a\| : a \in A\}$. The Hausdorff distance between two compact sets A and C in $\mathbb{R}^d$ is defined by

$$d_{\mathrm{H}}(A, C) = \max\left\{\sup_{a \in A} d(a, C), \sup_{c \in C} d(c, A)\right\} = \min\{\varepsilon > 0 : A \subset B(C, \varepsilon), C \subset B(A, \varepsilon)\}.$$

In many situations the boundary of G is a $(d - 1)$-dimensional manifold. The Hausdorff measure is often used in fractal geometry to determine the content of subsets whose dimension is lower than d; see, e.g. [7, pp. 54–74]. However, the Hausdorff measure is difficult to estimate when G is not completely known. Hence, we propose to measure the area of $T = \partial G$ in a simpler way (equivalent to the Hausdorff measure if T is smooth). We define the surface area of G by

$$L_0 \equiv L_0(G) = \lim_{\varepsilon \downarrow 0} L(\varepsilon), \tag{1}$$

provided that the limit exists and is finite, where

$$L(\varepsilon) = \frac{\mu(B(T, \varepsilon))}{2\varepsilon}. \tag{2}$$

We assume that G and $\overline{G^c}$ are both r-convex. A closed subset $A \subset \mathbb{R}^d$ is said to be r-convex if $A = C_r(A)$, where

$$C_r(A) = \bigcap_{\{\mathring{B}(x, r) : \, \mathring{B}(x, r) \cap A = \varnothing\}} (\mathring{B}(x, r))^c$$

is the smallest r-convex set that contains A. A set is r-convex if any point of the complement is contained in an open ball of radius r which does not meet the set. Thus, $C_r(A)$ will be close to $\bar{A}$ for small r and it will be close to the convex hull of A for large r; see, e.g. [8]. The r-convexity is also closely related to the dilation and erosion operations from mathematical morphology; see [10, pp. 43–59]. It can be easily seen that the r-convex hull of the set A is the morphological closing of A with respect to the structuring element $r\mathring{B}$, that is,

$$C_r(A) = (A \oplus r\mathring{B}) \ominus r\mathring{B},$$

where $\oplus$ and $\ominus$ respectively denote the Minkowski addition and Minkowski subtraction, defined by $A \oplus C := \{a+c \colon a \in A, \, c \in C\}$ and $A \ominus C := \{x \colon \{x\} \oplus C \subset A\}$, and $\lambda C := \{\lambda a \colon a \in A\}$ for any $A, C \subset \mathbb{R}^d$ and $\lambda \in \mathbb{R}$.

The sets G such that G and $\overline{G^c}$ are r-convex were characterized in [11]. It could be useful to state this characterization here. First, we need to define a very intuitive geometric concept. A ball of radius λ is said to roll freely in a closed set A if, for each point $a \in \partial A$, there exists $x \in \mathbb{R}^d$ such that $a \in B(x, \lambda) \subset A$.

Theorem 1. ([11].) *Let $G \subset \mathbb{R}^d$ be a nonempty compact set, and let $r > 0$. Then, the following conditions are equivalent.*

(i) *The conditions*

$$G = (G \ominus \lambda B) \oplus \lambda B, \qquad 0 \le \lambda \le r,$$
$$G = (G \oplus \lambda B) \ominus \lambda B, \qquad 0 \le \lambda < r,$$

    hold.

(ii) *G and $\overline{G^c}$ are r-convex and $\mathrm{int}(G_i) \ne \varnothing$ for each path-connected component $G_i \subset G$.*

(iii) *A ball of radius λ rolls freely inside G and $\overline{G^c}$ for all $0 \le \lambda \le r$.*

(iv) *∂G is a $(d-1)$-dimensional C^1 manifold with the outward pointing unit normal vector $\eta(x)$ at $x \in \partial G$, satisfying the Lipschitz condition*

$$\|\eta(x) - \eta(t)\| \le \frac{1}{r}\|x - t\| \quad \text{for all } x, t \in \partial G.$$

3. The sampling model and the estimator

Let G satisfy one of the conditions of Theorem 1. In [4] it was proved that L_0 is well defined for these sets. As in that paper, we assume that the sampling information is given by independent and identically distributed observations $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$ of a random variable (Z, δ), where Z is uniformly distributed on the unit square $[0, 1]^d$ and $\delta = \mathbf{1}_{\{Z \in G\}}$.

Let us denote by P_X and P_Y the conditional distributions of the observations in G and in G^c, respectively, that is, the uniform distributions on G and $R = [0, 1]^d \setminus \mathrm{int}(G)$, respectively.

Let $\{\varepsilon_n\}$ be a deterministic sequence of positive numbers which converges to 0 as n tends to ∞. We propose to estimate L_0 by means of

$$L_n = \frac{\mu(T_n)}{2\varepsilon_n}, \tag{3}$$

where T_n is an estimator of $B(T, \varepsilon_n)$. Based on the fact that

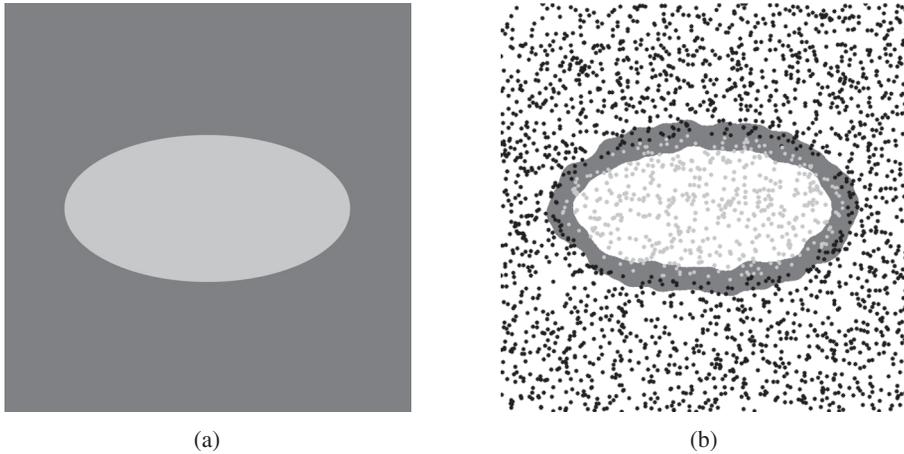$$B(T, \varepsilon_n) = B(G, \varepsilon_n) \cap B(R, \varepsilon_n),$$

FIGURE 1: (a) G and R in light and dark gray, respectively. (b) Uniform sample on $[0, 1]^2$ of size $n = 2500$, $\mathcal{X}_n$ (*light-gray dots*), $\mathcal{Y}_n$ (*black dots*), and T_n (*dark-gray region*) for $\varepsilon_n = 0.04$.

it is possible to construct an estimator of $B(T, \varepsilon_n)$ from estimators of the sets G and R. Since G and R are both r-convex, it is reasonable to estimate them using the r-convex hull of $\mathcal{X}_n = \{Z_i : \delta_i = 1\}$ and $\mathcal{Y}_n = \{Z_i : \delta_i = 0\}$. Thus, we propose the following estimator:

$$T_n = B(G_n, \varepsilon_n) \cap B(R_n, \varepsilon_n), \tag{4}$$

where

$$G_n = C_r(\mathcal{X}_n) \quad \text{and} \quad R_n = C_r(\mathcal{Y}_n).$$

An example of the proposed estimator is shown in Figure 1.

Remark 1. In [8] the r-convex hull of a sample is proposed as an estimator of the support of a distribution. Its asymptotic behavior is analyzed when the support is under the conditions of Theorem 1. The situation in this work is quite similar. In our case $\mathcal{X}_n$ and $\mathcal{Y}_n$ are samples (of random size) from the distributions P_X and P_Y with supports G and R, respectively. Therefore, many of the results in [8] can be adapted to this new situation, although it is necessary to do some adjustments since now the samples have random size and the boundary of R is not only T, it also includes the boundary of $[0, 1]^d$. This set is not under the assumptions of Theorem 1; see the rolling condition, Theorem 1(iii).

4. Theoretical results

Theorem 2 gives the almost sure rate of convergence of L_n to L_0 and represents the main result of this work.

Theorem 2. *Let $G \subset (0, 1)^d$ be a set under the conditions of Theorem 1. Then, with probability 1,*

$$\inf_{\varepsilon_n} |L_n - L_0| = O\left(\frac{\log n}{n}\right)^{1/(d+1)}.$$

Remark 2. In [4], a similar estimator to the one defined in this work is proposed. They considered the estimator given in (4) with $G_n = \mathcal{X}_n$ and $R_n = \mathcal{Y}_n$ and proved its universal

consistency (provided that $L_0(G)$ existed). However, the convergence rate for the estimator of L_0 based on this empirical approximation of $B(T, \varepsilon_n)$ is $n^{-1/2d}$, which is worse than the rate attained with the estimator proposed in this work. The main reason for this improvement is that smoothing the samples $\mathcal{X}_n$ and $\mathcal{Y}_n$ allows us to choose a smaller radius ε_n of order $(\log n/n)^{1/(d+1)}$. In [4] the optimal ε_n had to be of order $n^{-1/2d}$. An important practical problem is the construction of a data-driven selection method for the smoothing parameter ε_n. In that sense, a simulation study in [4] shows a considerable stability of the estimator with respect to the values of ε_n. We have obtained the same conclusion for L_n in a preliminary simulation study.

4.1. Proof of Theorem 2

We will follow the ideas of the proof of Theorem 3 of [8]. The proof is based on Propositions 1, 2, and 3, below. Proposition 1 establishes that if $T \subset B(Z_n^{\mathcal{X}}, 2\rho_n) \cap B(Z_n^{\mathcal{Y}}, 2\rho_n)$, where $Z_n^{\mathcal{X}} = \{Z_i \in \mathcal{X}_n : d(Z_i, T) \leq \rho_n^2\}$ and $Z_n^{\mathcal{Y}} = \{Z_i \in \mathcal{Y}_n : d(Z_i, T) \leq \rho_n^2\}$, then $B(T, \varepsilon_n) \setminus T_n$ is contained in the disc $D_n = B(T, \varepsilon_n) \setminus B(T, \varepsilon_n - K\rho_n^2)$ for large enough K. Proposition 2 relies on the fact that $\mu(D_n) = O(\rho_n^2)$ to find a bound for $|L_n - L_0|$ depending only on ε_n and ρ_n. Finally, in Proposition 3 we determine the order of ρ_n for which, with probability 1, we have $T \subset B(Z_n^{\mathcal{X}}, 2\rho_n) \cap B(Z_n^{\mathcal{Y}}, 2\rho_n)$ for large enough n, that is, ρ_n satisfies

$$P(T \subset B(Z_n^{\mathcal{X}}, 2\rho_n) \cap B(Z_n^{\mathcal{Y}}, 2\rho_n) \text{ eventually}) = 1.$$

Theorem 2 is a straightforward consequence of these three results.

Proposition 1. *Let G be a set under the conditions of Theorem 2. Then the following results hold.*

(i) *With probability 1, $T_n \subset B(T, \varepsilon_n)$.*

(ii) *Let us assume that $\rho_n \to 0$ satisfies $\rho_n^2 \varepsilon_n^{-1} \to 0$ and that*

$$P(T \subset B(Z_n^{\mathcal{X}}, 2\rho_n) \cap B(Z_n^{\mathcal{Y}}, 2\rho_n) \text{ eventually}) = 1,$$

where $Z_n^{\mathcal{X}} = \{Z_i \in \mathcal{X}_n : d(Z_i, T) \leq \rho_n^2\}$ and $Z_n^{\mathcal{Y}} = \{Z_i \in \mathcal{Y}_n : d(Z_i, T) \leq \rho_n^2\}$. Then, if $K \geq \max(2, 8/r)$, we have

$$P(B(T, \varepsilon_n - K\rho_n^2) \subset T_n \text{ eventually}) = 1.$$

Proof. Under the conditions of the proposition, G and R are r-convex. Since, with probability 1, $\mathcal{X}_n \subset G$ and $\mathcal{Y}_n \subset R$,

$$G_n = C_r(\mathcal{X}_n) \subset C_r(G) = G, \qquad R_n = C_r(\mathcal{Y}_n) \subset C_r(R) = R.$$

Thus, with probability 1,

$$T_n = B(G_n, \varepsilon_n) \cap B(R_n, \varepsilon_n) \subset B(G, \varepsilon_n) \cap B(R, \varepsilon_n) = B(T, \varepsilon_n),$$

which concludes the proof of (i). On the other hand, the proof of (ii) will be finished if we show that

$$P(T \subset B(G_n, K\rho_n^2) \cap B(R_n, K\rho_n^2) \text{ eventually}) = 1, \tag{5}$$

since if $T \subset B(G_n, K\rho_n^2) \cap B(R_n, K\rho_n^2)$ and $\varepsilon_n > K\rho_n^2$ then

$$
\begin{aligned}
B(T, \varepsilon_n - K\rho_n^2) &\subset B(B(G_n, K\rho_n^2) \cap B(R_n, K\rho_n^2), \varepsilon_n - K\rho_n^2) \\
&\subset B(G_n, \varepsilon_n) \cap B(R_n, \varepsilon_n) \\
&= T_n.
\end{aligned}
$$

In order to prove (5), it suffices to show that, with probability 1, for large enough n,

$$
x_G = x - K\rho_n^2 \eta(x) \in G_n \quad \text{and} \quad x_R = x + K\rho_n^2 \eta(x) \in R_n \quad \text{for } x \in T, \tag{6}
$$

where $\eta(x)$ is the outward pointing unit normal vector at x (see Theorem 1(iv)).

To prove (6), we have to show that x_G cannot be contained in an open ball of radius r which does not meet the sample $\mathcal{X}_n$. In the same manner we have to prove that x_R cannot be contained in an open ball of radius r which does not meet the sample $\mathcal{Y}_n$. The situation in which the center of the ball is close to T is analyzed in Lemma 2, below. This lemma yields the result for x_G. For x_R, we have also to analyze the situation in which the center of the ball is far from T. This case is studied in Lemma 3, below. Finally, in Lemma 4, below, both results are used to establish the precise conditions under which (6) is satisfied. The proposition is a consequence of this result. We begin with a geometric lemma, needed to prove Lemma 2.

Lemma 1. *Let G be a set under the conditions of Theorem 1, and let $y \in \mathbb{R}^d$ such that $d(y, T) = r - \delta$, where $0 \leq \delta \leq r$. Then, for all $x \in \mathbb{R}^d$ with $d(x, T) \leq \delta/2$ and $\|x - y\| \geq r$, we have*

$$
\|x - P_T y\| \geq \sqrt{\frac{r\delta}{2}},
$$

where $P_T y$ is the metric projection of y onto T.

Proof. This lemma is similar to Lemma 1 of [8] and its proof is almost identical. Let $y \in \mathbb{R}^d$ be a point such that $d(y, T) = r - \delta$, where $0 \leq \delta \leq r$. We denote by $P_T y$ the metric projection of y onto T. Let η be the outward pointing unit normal vector at $P_T y$.

First, we assume that $y \in G$. Then, it is easy to prove that $y = P_T y - (r - \delta)\eta$. Let $t = P_T y + r\eta$. Then, for $x \in \mathbb{R}^d$ with $d(x, T) \leq \delta/2$ and $\|x - y\| \geq r$,

$$
r^2 \leq \|x - y\|^2 = \|x - P_T y + (r - \delta)\eta\|^2 = \|x - P_T y\|^2 + (r - \delta)^2 + 2(r - \delta)\langle x - P_T y, \eta \rangle,
$$

$$
\left(r - \frac{\delta}{2}\right)^2 \leq \|x - t\|^2 = \|x - P_T y - r\eta\|^2 = \|x - P_T y\|^2 + r^2 - 2r\langle x - P_T y, \eta \rangle.
$$

The second inequality is a consequence of $d(x, T) \leq \delta/2$ and $d(t, T) = r$. Then

$$
\|x - P_T y\|^2 + 2(r - \delta)\langle x - P_T y, \eta \rangle \geq 2r\delta - \delta^2,
$$

$$
\|x - P_T y\|^2 - 2r\langle x - P_T y, \eta \rangle \geq -r\delta + \frac{\delta^2}{4}.
$$

Multiplying the first inequality by r and the second inequality by $(r - \delta)$ and adding, we have

$$
\|x - P_T y\|^2 \geq \frac{2r^2\delta - r\delta^2 - (r - \delta)r\delta + (r - \delta)(\delta^2/4)}{2r - \delta} = \frac{r^2\delta + (r - \delta)(\delta^2/4)}{2r - \delta} \geq \frac{r\delta}{2},
$$

where the last inequality is a consequence of $0 \leq \delta \leq r$.

For $y \in G^c$, we can apply the previous result to $\overline{G^c}$. In this case $y \in \overline{G^c}$ and T is also the boundary of $\overline{G^c}$.

Lemma 2. *Let G be a set under the conditions of Theorem 2, and let*

$$T \subset B(Z_n^{\mathcal{X}}, 2\rho_n) \cap B(Z_n^{\mathcal{Y}}, 2\rho_n),$$

where $Z_n^{\mathcal{X}} = \{Z_i \in \mathcal{X}_n : d(Z_i, T) \leq \rho_n^2\}$ and $Z_n^{\mathcal{Y}} = \{Z_i \in \mathcal{Y}_n : d(Z_i, T) \leq \rho_n^2\}$. Then, for all $y \in \mathbb{R}^d$ such that $d(y, T) = r - \delta$ with $\max(2, 8/r)\rho_n^2 < \delta \leq r$,

$$\mathring{B}(y, r) \cap \mathcal{X}_n \neq \varnothing \quad and \quad \mathring{B}(y, r) \cap \mathcal{Y}_n \neq \varnothing.$$

Proof. Let $y \in \mathbb{R}^d$ be a point such that $d(y, T) = r - \delta$ with $\max(2, 8/r)\rho_n^2 < \delta \leq r$. We denote by $P_T y$ the metric projection of y onto T. Since $T \subset B(Z_n^{\mathcal{X}}, 2\rho_n)$, there exists $z_x \in Z_n^{\mathcal{X}}$ such that $\|z_x - P_T y\| \leq 2\rho_n$. Furthermore, $d(z_x, T) \leq \rho_n^2 < \delta/2$. If $\|z_x - y\| \geq r$ then Lemma 1 yields

$$\|z_x - P_T y\| \geq \sqrt{\frac{r\delta}{2}} > 2\rho_n,$$

which leads to a contradiction. The last inequality is a consequence of $\delta > 8\rho_n^2/r$. Therefore, $\|z_x - y\| < r$ and $\mathring{B}(y, r) \cap \mathcal{X}_n \neq \varnothing$. In the same manner it can be proved that $\mathring{B}(y, r) \cap \mathcal{Y}_n \neq \varnothing$.

Before stating Lemma 3, it is necessary to introduce some notation. Since we are assuming that $G \subset (0, 1)^d$, for all $x \in G$,

$$d(x, \mathbb{R}^d \setminus (0, 1)^d) > 0.$$

The function $d(\cdot, \mathbb{R}^d \setminus (0, 1)^d)$ is continuous and, therefore, it reaches its minimum in the compact set G. Let us denote by e the minimum, that is,

$$e = \min_{x \in G} d(x, \mathbb{R}^d \setminus (0, 1)^d) > 0. \tag{7}$$

Note that $B(G, e) \subset [0, 1]^d$.

Lemma 3. *Let $x \in \mathbb{R}^d$ be a point such that $0 \leq d(x, G) \leq e/2$, and let $y \notin [0, 1]^d$ satisfy $x \in \mathring{B}(y, r)$. Then there exists $z_0 \in R$ for which $B(z_0, e/4) \subset \mathring{B}(y, r)$.*

Proof. The function

$$d(\lambda) = d(\lambda x + (1 - \lambda)y, G), \qquad 0 \leq \lambda \leq 1,$$

is continuous. Since $y \notin [0, 1]^d$, we have $d(0) = d(y, G) > e$. Furthermore, $d(1) = d(x, G) \leq e/2$. Bolzano's theorem establishes that there exists z_0 in the segment joining x and y such that $d(z_0, G) = 3e/4$. Moreover, $z_0 \in R$, since $z_0 \in B(G, e) \subset [0, 1]^d$ and $z_0 \notin G$.

Let us now prove that $B(z_0, e/4) \subset \mathring{B}(y, r)$. Let $z \in B(z_0, e/4)$. We have

$$\|z - y\| \leq \|z - z_0\| + \|z_0 - y\| \leq \frac{e}{4} + \|z_0 - y\|.$$

Since z_0 is in the segment joining x and y, $\|z_0 - y\| = \|x - y\| - \|x - z_0\|$. From $d(z_0, G) = 3e/4$ and $d(x, G) \leq e/2$, it follows that $\|x - z_0\| \geq e/4$ and, therefore,

$$\|z_0 - y\| = \|x - y\| - \|x - z_0\| < r - \frac{e}{4}.$$

Thus,

$$\|z - y\| < \frac{e}{4} + r - \frac{e}{4} = r.$$

Lemma 4. *Let us assume that $T \subset B(Z_n^{\mathcal{X}}, 2\rho_n) \cap B(Z_n^{\mathcal{Y}}, 2\rho_n)$ and $K\rho_n^2 < \min(e/2, r)$ with $K \geq \max(2, 8/r)$. Let us also assume that $d_{\mathrm{H}}(\mathcal{X}_n, G) < r$ and $d_{\mathrm{H}}(\mathcal{Y}_n, R) < \min(e/4, r)$. Then, for all $x \in T$,*

$$x - K\rho_n^2 \eta(x) \in G_n \quad and \quad x + K\rho_n^2 \eta(x) \in R_n,$$

where $\eta(x)$ is the outward pointing unit normal vector at x.

Proof. Let $x \in T$, and let $x_G = x - K\rho_n^2\eta(x)$. The point x_G belongs to G_n if any open ball of radius r that contains the point x_G meets the sample $\mathcal{X}_n$. Thus, let $y \in \mathbb{R}^d$ be a point such that $x_G \in \mathring{B}(y, r)$. We want to show that $\mathring{B}(y, r) \cap \mathcal{X}_n$ is not empty. This is straightforward when $y \in G$, since $d_{\mathrm{H}}(\mathcal{X}_n, G) < r$. Now, let us assume that $y \in G^c$. Since $x_G \in \mathring{B}(y, r) \cap G$ ($K\rho_n^2 < r$), then $d(y, T) = r - \delta$, where $\delta > K\rho_n^2 \geq \max(2, 8/r)\rho_n^2$. By Lemma 2 we have $\mathring{B}(y, r) \cap \mathcal{X}_n \neq \varnothing$.

Now, let $x_R = x + K\rho_n^2\eta(x)$. As before, in order to prove that x_R belongs to R_n, we need to show that $\mathring{B}(y, r) \cap \mathcal{Y}_n$ is not empty for any $y \in \mathbb{R}^d$ such that $x_R \in \mathring{B}(y, r)$. Again, this is straightforward when $y \in R$, since $d_{\mathrm{H}}(\mathcal{Y}_n, R) < r$ by assumption. Now let us assume that $y \notin R$. There are two possibilities: $y \in G$ or $y \notin [0, 1]^d$. For $y \in G$, as $x_R \in \mathring{B}(y, r) \cap G^c$ ($K\rho_n^2 < r$), we have $d(y, T) = r - \delta$ with $\delta > K\rho_n^2 \geq \max(2, 8/r)\rho_n^2$. Lemma 2 implies that $\mathring{B}(y, r) \cap \mathcal{Y}_n \neq \varnothing$. Finally, if $y \notin [0, 1]^d$, by the definition of x_R, we have $d(x_R, G) = K\rho_n^2 < e/2$. Then Lemma 3 establishes that there exists $z_0 \in R$ such that $B(z_0, e/4) \subset \mathring{B}(y, r)$. Since $d_{\mathrm{H}}(\mathcal{Y}_n, R) < e/4$, we have $B(z_0, e/4) \cap \mathcal{Y}_n \neq \varnothing$. Thus, we have $\mathring{B}(y, r) \cap \mathcal{Y}_n \neq \varnothing$.

The proof of Proposition 1 is now complete since the conditions of Lemma 4 are satisfied with probability 1 for large enough n.

Now, we can prove Proposition 2, below, which gives a bound for the distance $|L_n - L_0|$ by splitting it into a *bias* term $|L(\varepsilon_n) - L_0|$ and a *variance* term $|L_n - L(\varepsilon_n)|$. Recall that L_0, $L(\varepsilon_n)$, and L_n were defined in (1), (2), and (3), respectively.

Proposition 2. *Under the conditions of Proposition 1, we have, with probability 1,*

$$|L_n - L_0| \leq |L(\varepsilon_n) - L_0| + O\left(\frac{\rho_n^2}{\varepsilon_n}\right) = O(\varepsilon_n) + O\left(\frac{\rho_n^2}{\varepsilon_n}\right)$$

and

$$\inf_{\varepsilon_n} |L_n - L_0| = O(\rho_n).$$

Proof. We have

$$|L_n - L_0| \leq |L_n - L(\varepsilon_n)| + |L(\varepsilon_n) - L_0|.$$

In [4] it was proved that $|L(\varepsilon_n) - L_0| = O(\varepsilon_n)$. On the other hand, Proposition 1 yields

$$\mathrm{P}(B(T, \varepsilon_n - K\rho_n^2) \subset T_n \subset B(T, \varepsilon_n) \text{ eventually}) = 1.$$

Then, with probability 1, for large enough n,

$$|L_n - L(\varepsilon_n)| = \frac{\mu(B(T, \varepsilon_n))}{2\varepsilon_n} - \frac{\mu(T_n)}{2\varepsilon_n} \leq \frac{\mu(B(T, \varepsilon_n)) - \mu(B(T, \varepsilon_n - K\rho_n^2))}{2\varepsilon_n}.$$

In the following lemma the convergence rate of the last term in the previous inequality is determined.

Lemma 5. *Assume that $F(\varepsilon) = \mu(B(T, \varepsilon))$ is differentiable in a neighborhood of 0 and that the derivative F' is continuous at 0. Then*

$$\lim_{n \to \infty} \frac{\mu(D_n)}{2K\rho_n^2} = L_0,$$

where $D_n = B(T, \varepsilon_n) \setminus B(T, \varepsilon_n - K\rho_n^2)$.

Remark 3. In Lemma 5 we assume that the function $F(\varepsilon) = \mu(B(T, \varepsilon))$ is smooth in a neighborhood of 0. This holds for the boundary T of a set G under the conditions of Theorem 1. In that case it can be proved that T satisfies the conditions of Theorem 5.6 of [6]. This result ensures that $F(\varepsilon)$ coincides locally for $\varepsilon \in (0, r)$ with a polynomial of degree at most d. Therefore, Lemma 5 suffices for our purposes.

Proof of Lemma 5. For large enough n (since $\varepsilon_n, \rho_n \to 0$), we have

$$
\begin{aligned}
\frac{\mu(D_n)}{2K\rho_n^2} &= \frac{\mu(B(T, \varepsilon_n)) - \mu(B(T, \varepsilon_n - K\rho_n^2))}{2K\rho_n^2} \\
&= \frac{F(\varepsilon_n) - F(\varepsilon_n - K\rho_n^2)}{2K\rho_n^2} \\
&= \frac{F'(\xi_n)K\rho_n^2}{2K\rho_n^2} \\
&= \frac{F'(\xi_n)}{2},
\end{aligned}
$$

where ξ_n is a point in the interval $(\varepsilon_n - K\rho_n^2, \varepsilon_n)$ (by the mean value theorem). Since F' is continuous at 0,

$$\lim_{n \to \infty} \frac{\mu(D_n)}{2K\rho_n^2} = \frac{F'(0)}{2} = L_0,$$

where the last equality is a consequence of (1).

By Lemma 5, with probability 1,

$$|L_n - L(\varepsilon_n)| = O\left(\frac{\rho_n^2}{\varepsilon_n}\right).$$

Therefore, with probability 1,

$$|L_n - L_0| \le |L(\varepsilon_n) - L_0| + O\left(\frac{\rho_n^2}{\varepsilon_n}\right) = O(\varepsilon_n) + O\left(\frac{\rho_n^2}{\varepsilon_n}\right).$$

To prove the last part of Proposition 2, it suffices to take $\varepsilon_n = \rho_n$.

As we mentioned at the beginning of Subsection 4.1, in the following proposition we determine the rate for ρ_n which guarantees that, with probability 1, $T \subset B(Z_n^{\mathcal{X}}, 2\rho_n) \cap B(Z_n^{\mathcal{Y}}, 2\rho_n)$ for large enough n.

Proposition 3. *If $c > 0$ is large enough then*

$$\mathrm{P}(T \subset B(Z_n^{\mathcal{X}}, 2\rho_n) \cap B(Z_n^{\mathcal{Y}}, 2\rho_n) \text{ eventually}) = 1,$$

where

$$\rho_n = \left(\frac{c \log n}{n} \right)^{1/(d+1)},$$

$Z_n^{\mathcal{X}} = \{Z_i \in \mathcal{X}_n \colon d(Z_i, T) \le \rho_n^2\}$, *and* $Z_n^{\mathcal{Y}} = \{Z_i \in \mathcal{Y}_n \colon d(Z_i, T) \le \rho_n^2\}$.

Proof. Theorem 1 of [5] establishes that, for $\rho_n > 0$,

$$\mathrm{P}(T \not\subseteq B(Z_n^{\mathcal{U}}, 2\rho_n)) \le \rho_n^{-d} \Pi(G, Z_n^{\mathcal{U}}, \rho_n), \qquad \mathcal{U} = \mathcal{X}, \mathcal{Y},$$

where $\Pi(G, Z_n^{\mathcal{U}}, \rho_n) = \sup_{x \in T} \mathrm{P}(B(x, \rho_n) \cap Z_n^{\mathcal{U}} = \varnothing)$. Therefore, by the Borel–Cantelli lemma, it remains to prove that

$$\sum_{n=1}^{\infty} \rho_n^{-d} \Pi(G, Z_n^{\mathcal{U}}, \rho_n) < \infty, \qquad \mathcal{U} = \mathcal{X}, \mathcal{Y}. \tag{8}$$

Let $x \in T$. Since Z is uniformly distributed on $[0, 1]^d$, we have, for $\rho_n^2 < e$ (see (7)),

$$
\begin{aligned}
\mathrm{P}(B(x, \rho_n) \cap Z_n^{\mathcal{X}} = \varnothing) &= \mathrm{P}(Z_i \notin B(x, \rho_n) \cap B(T, \rho_n^2) \cap G, \ i = 1, \dots, n) \\
&= (1 - \mu(B(x, \rho_n) \cap B(T, \rho_n^2) \cap G))^n \\
&\le \exp(-n\mu(B(x, \rho_n) \cap B(T, \rho_n^2) \cap G)).
\end{aligned}
$$

Likewise,

$$\mathrm{P}(B(x, \rho_n) \cap Z_n^{\mathcal{Y}} = \varnothing) \le \exp(-n\mu(B(x, \rho_n) \cap B(T, \rho_n^2) \cap R)).$$

The following lemma is proved in [8]. It gives a lower bound for $\mu(B(x, \rho_n) \cap B(T, \rho_n^2) \cap G)$ and $\mu(B(x, \rho_n) \cap B(T, \rho_n^2) \cap R)$ for large enough n.

Lemma 6. *If G is under the conditions of Theorem 1 then there exist constants $\alpha, \beta > 0$ such that, for all $\varepsilon \in [0, \beta]$ and for all $x \in T$,*

$$\mu(B(x, \varepsilon) \cap B(T, \varepsilon^2) \cap G) \ge \alpha \varepsilon^{d+1} \quad and \quad \mu(B(x, \varepsilon) \cap B(T, \varepsilon^2) \cap \overline{G^c}) \ge \alpha \varepsilon^{d+1}.$$

It is easy to prove that if ρ_n is small enough (for example, $\rho_n^2 < e$) then

$$\mu(B(x, \rho_n) \cap B(T, \rho_n^2) \cap R) = \mu(B(x, \rho_n) \cap B(T, \rho_n^2) \cap \overline{G^c}),$$

and, therefore, by Lemma 6,

$$\mathrm{P}(B(x, \rho_n) \cap Z_n^{\mathcal{U}} = \varnothing) \le \exp(-n\alpha\rho_n^{d+1}), \qquad \mathcal{U} = \mathcal{X}, \mathcal{Y}.$$

It is not hard to prove that, for large enough c, (8) is satisfied. Thus, the proof of the proposition is complete.

Theorem 2 is a straightforward consequence of Propositions 1, 2, and 3.

Acknowledgements

References

[1] BADDELEY, A. AND JENSEN, E. B. V. (2005). *Stereology for Statisticians*. Chapman & Hall/CRC, Boca Raton, FL.

[2] BRÄKER, H. AND HSING, T. (1998). On the area and perimeter of a random hull in a bounded convex set. *Prob. Theory Relat. Fields* **111,** 517–550.

[3] CUEVAS, A. AND RODRÍGUEZ-CASAL, A. (2003). Set estimation: an overview and some recent developments. In *Recent Advances and Trends in Nonparametric Statistics*, eds M. G. Akritas and D. N. Politis, Elsevier, Amsterdam, pp. 251–264.

[4] CUEVAS, A., FRAIMAN, R. AND RODRÍGUEZ-CASAL, A. (2007). A nonparametric approach to the estimation of lengths and surface areas. *Ann. Statist.* **35,** 1031–1051.

[5] DÜMBGEN, L. AND WALTHER, G. (1996). Rates of convergence for random approximations of convex sets. *Adv. Appl. Prob.* **28,** 384–393.

[6] FEDERER, H. (1959). Curvature measures. *Trans. Amer. Math. Soc.* **93,** 418–491.

[7] MATTILA, P. (1995). *Geometry of Sets and Measures in Euclidean Spaces*. Cambridge University Press.

[8] RODRÍGUEZ-CASAL, A. (2007). Set estimation under convexity type assumptions. *Ann. Inst. H. Poincaré Prob. Statist.* **43,** 763–774.

[9] SCHNEIDER, R. (1988). Random approximation of convex sets. *J. Microscopy* **151,** 211–227.

[10] SERRA, J. (1982). *Image Analysis and Mathematical Morphology*. Academic Press, London.

[11] WALTHER, G. (1999). On a generalization of Blaschke's rolling theorem and the smoothing of surfaces. *Math. Methods Appl. Sci.* **22,** 301–316.